ADVANCES IN MATHEMATICS 40, 176–200 (1981)

An Example in Axiomatic Set Theory

E. M. KLEINBERG

Department of Mathematics, State University of New York, Buffalo, New York 14214

PREFACE

This is meant to be an expository paper. Its goal is to present several recent major developments in set theory. We shall be writing for a general audience, and despite the technical limitations this imposes on us, we will prove several deep theorems, a number of which have traditional mathematical interest quite apart from any considerations of logic. Our main hope is to do justice to the power and beauty of the techniques involved and, in general, to capture some of the excitement surrounding the field.

Strictly speaking, the only prerequisite for an understanding of what follows is a sound foundation in undergraduate mathematics.

The only notion specifically from set theory with which we shall assume familiarity is that of "ordinal number." For those who are unfamiliar with this concept perhaps this will help: the nonnegative integers 0, 1, 2... are the finite ordinal numbers. To get all the ordinal numbers one should simply continue counting. The first infinite ordinal is called ω, the next $\omega + 1$, then $\omega + 2$, $\omega + 3$, $\omega + 4$, $\omega + 5$,... eventually $\omega + \omega$, $\omega + \omega + 1$, $\omega + \omega + 2$,... $\omega + \omega + \omega$,... $\omega \cdot \omega$, $\omega \cdot \omega + 1$,... and so forth. For our purposes, the two key facts about the ordinals are these: (1) the collection of ordinals forms a well-ordering, that is, the ordinals are linearly ordered and every nonempty set of ordinals has a least member; and (2) there exist "arbitrarily many" ordinals, that is, given any set there exists a set of ordinals which is so large that it cannot be mapped one-to-one into the set.

One more word before we begin: except in a few cases of theorems commonly referred to via their discoverer's name, we will not attribute credit for theorems, or proofs, in the body of the text. Such acknowledgements will be given later, just prior to the list of references.

1

The field of mathematical logic has not been particularly successful in discovering the logic of mathematics. Indeed, aside from seemingly super-

0001-8708/81/050176–25$05.00/0

ficial formalisms, a working theory of mathematical practice simply does not exist. What has, from a foundational point of view, been the major success of logic is the description of the limits of various aspects of mathematics. Results along these lines range from the abstract theorems of Gödel stating that, associated with any given notion of mathematical proof (which is consistent and deductive in nature), there will be statements of mathematics which are neither provable nor disprovable, to the more recent concrete independence theorems in set theory showing that particular well-known mathematical statements are neither provable nor disprovable.

Let us be more specific concerning these independence results. During the early portion of the 20th century, a formal axiomatic theory known as Zermelo–Fraenkel set theory, or ZF, was developed to serve as a foundation for mathematics. Its motivation was as follows: so far as anyone had been able to determine, all standard mathematical notions were representable in the language of sets. The theory ZF was created by simply listing a collection of natural axioms about sets (for example: "two sets are equal if they have the same members"; *or* "given sets x and y there exists a set whose members are precisely the sets x and y") along with the standard rules of logical inference. At any rate, with the inclusion of the axiom of choice, the resulting theory, ZFC, seemed (and still seems) to be general enough to contain virtually *all* of mathematics. This is not to imply that mathematicians have the axioms or rules of ZFC in mind when they work, or that they even know these axioms and rules. We are only saying that to the extent to which it has been possible to judge, if one took the trouble to translate mathematics into the language of sets, the theory of ZFC would be broad enough to prove any theorem tried. The virtue of this is that in ZFC we have a very *simple and easy to use representation of the range of mathematics* or, more precisely, of the range of what mathematicians currently think of as mathematics. And such a representation is almost essential if one is to prove results about the limits of this range. What we have in mind is this: occasionally a mathematician must make special assumptions in proving a desired theorem and he naturally questions whether or not the special assumptions themselves can be proved. The continuum hypothesis, for example, is sometimes used in proving theorems in mathematics[1]—is the continuum hypothesis itself provable? Well, if we consider the mathematically concrete theory ZFC as embodying the realm of acceptable proof, the answer is a definitive no. This, the famous independence result of Cohen, in effect puts a very specific bound on mathematics in its present form.

As with Gödel's celebrated results of thirty years earlier, much misunderstanding followed the announcement of Cohen's work. The continuum

[1] For example, the Von Neumann–Stone lifting theorem.

hypothesis is still either true or false—nothing can change that. It is just that ZFC is not sufficiently rich to decide which.

Fortunately, mathematics tends to expand. This happens both through the discovery of new concepts and techniques, and through the introduction, increased use, and general acceptance of new principles. Sometimes, as evidenced by the evolution of nonstandard analysis, work which initially seems dependent upon new principles or axioms turns out to be accessible without them.

In this paper we examine a technique of D. A. Martin which grew out of work associated with various new mathematical axioms. Two examples of proofs involving Martin's technique will be given. One of these is dependent upon such a new axiom for set theory and the other is not. Although we do not wish to discuss here the virtue of pursuing a systematic study of new axioms, it is important to note the following: when proving theorems from even the most exotic, unnatural, or implausible axioms there is often associated a realistic hope of establishing the very same results *without* the special axioms. For such axioms or hypotheses tend to present an atmosphere of conceptual or notational simplicity from which various techniques or methods naturally develop. A typical scenario may go as follows: one first proves a theorem using a special axiom and then carefully examines his proof, however complicated this analysis may become, to see just how much of the power of the axiom was really needed to establish the result. Hopefully he will find that for the context at hand some modification of the axiom *which is provable outright in ZF* is all that is needed. In this case he would have his result as a theorem of ZF by a method he never would have discovered directly. This precise thing has happened and we shall see an example of it later.

It would be unrealistic at this point to discuss Martin's method per se. However, in order to tantalize the reader, we might mention the following: although our main theorems are facts about "small" sets such as the set of reals, their proofs draw heavily upon the existence of extremely large sets, some much larger than

$$\text{the set of sets of } \underbrace{\cdots\cdots\cdots}_{\text{infinite}} \text{ of sets of reals.}$$

What possibly bearing can the existence of large sets have on proving facts about small sets? We don't wish to sound overly melodramatic but the solution is simple yet striking and shall unfold as we proceed.

Remark. Since not all readers of this paper will have the same backgrounds or interests, it might be helpful for us to make a number of remarks concerning the organization and structure of this paper, and in particular to point out which sections are very difficult or technical and to what extent they can be circumvented. As the title and introduction suggest, this paper can be thought of as an exposition of a number of related large cardinal properties and of related applications of these properties to results about sets of real numbers. In the section immediately following this general introduction, Section 2, we begin with a very elementary discussion of the notion of one abstract set being larger than another and lead naturally from this into a first approximation of the sorts of large cardinal properties to be considered later. Section 2 is thus an introduction to our subsequent work with large cardinal properties and it is short and easy. In Section 3 we introduce a notion related to sets of real numbers. It is to results about this notion that we will later apply our large cardinal properties. Section 3 is basically easy except for a theorem in it about Lebesgue measure, a theorem which should be of interest and enjoyment to those who are familiar with Lebesgue measure but which, since its purpose is mainly motivational, may be skipped by those who are not. Section 4 begins with an easy theorem but proceeds rapidly to a proof of the first main theorem of the paper. It is in this theorem that we present an application of the properties of Section 2 to proving an instance of the notion of Section 3, and although the proof is long, it contains interesting ideas and is not too difficult to follow. In Section 5 we present another application of large cardinal properties to results about sets of reals. Here, however, the application is far more sophisticated. In this section we will evolve from the sort of very simple large cardinal properties of Section 2, a much more refined sort of large cardinal property, and we shall apply the refined property to proving a much more general instance of the notion of Section 3. The proof associated with this application is difficult and may be skipped, but it is of extreme interest and is certainly worth the effort of understanding. Section 6 is the concluding section. It discusses in some detail the refined sort of large cardinal property introduced at the beginning of Section 5 and concludes with a number of remarks about large cardinal properties in general. This section is basically easy.

Hopefully this outline will help in reading what follows. Keep in mind that you can always return to it to regain any eroded perspective.

2

During the course of this paper we shall be dealing with various properties of "large" sets. Rather than pull these out of thin air let us attempt to arrive at them naturally.

The key property of "large" sets on which we plan to build is embodied in the so-called *pigeon-hole principle*: if a "large" set is partitioned into a "small" number of pieces, then one of the pieces must contain a "large" number of elements. But what do we *mean* by the words "large" and "small" as applied to arbitrary sets? These terms are of course relative. The generally accepted definition is that one set x is of *greater cardinality than* a set y (or, simply, is *larger than* y) if there does not exist a one to one mapping of x into y. Thus if x is of greater cardinality than y, then given any function f from x into y, f must fail to be one to one, that is, there must exist distinct members p and q of x such that $f(p) = f(q)$. As we can easily see, this is just an instance of the pigeon-hole principle, for a mapping from a set x into a set y is really the same as a partition of x into "y many" pieces the "rth piece" of the partition, for r a member of y, being simply $f^{-1}\{r\}$, the set of p in x such that $f(p) = r$.

But the pigeon-hole principle goes well beyond direct consequences of the definition of cardinality. Here is an example: "if an infinite set is partitioned into a finite number of pieces then one of the pieces will be infinite." By viewing partitions as functions as above we might write this "any function with infinite domain and finite range is constant on an infinite set."

The pigeon-hole principle is well known to mathematics. A prime example appears in the proof of the Bolzano–Weierstrass theorem which states that every infinite bounded set of reals has a limit point. For if S is an infinite bounded set of reals, the proof proceeds as follows: since S is bounded it intersects finitely many intervals of the form $[i, i+1)$, where i is an integer, and hence the members of S can be viewed as being partitioned into finitely many pieces according to which of these finitely many intervals they lie in. Thus by the pigeon-hole principle, since S is infinite, infinitely many members of S must all lie in some one interval, say $[i_0, i_0+1)$. (We may assume, without loss of generality, that $i_0 \geqslant 0$.) Let S_0 be the infinite subset of S contained in $[i_0, i_0+1)$. Then if we view $[i_0, i_0+1)$ as consisting of the ten subintervals $[i_0, i_0+1/10)$, $[i_0+1/10,\ i_0+2/10)$,..., $[i_0+9/10,\ i_0+1)$, another use of the pigeon-hole principle tells us that an infinite subset S_1, of S_0 is contained entirely within $[i_0+i_1/10,\ i_0+i_1+1/10)$ for some integer i_1, $0 \leqslant i_1 < 10$. Now dividing $[i_0+i_1/10,\ i_0+i_1+1/10)$ into ten successive subintervals and applying the pigeon-hole principle again we define i_2 and S_2, and eventually i_3 and S_3, i_4 and S_4, i_5 and S_5, and so forth ad infinitum. It is routine to now check that $i_0.\ i_1 i_2 i_3 i_4 i_5 \cdots$ is a decimal expansion of a desired limit point of our original set S.

The applications of the pigeon-hole principle and of related large cardinal axioms to be considered as we proceed will be entirely different from this one. For even though the theorems to be proved will still be about sets of real numbers, the sets being pigeon-holed will be sets which are much larger than the set of real numbers.

3

In this section we introduce the idea of a given set of reals being "determinate." It is notationally best to do this in terms of two-person games of infinite length, but as an initial (and necessarily vague) description, we might say that "determinateness" is an attempt to impose some sort of uniformity on a given set of reals in terms of a continuous function.

We begin by looking at the reals themselves. In particular we would like to set up some method for specifying real numbers which avoids the usual ambiguities inherent with decimals, that is, for example, the problem associated with the fact that 1.00000... and 0.99999... both represent the same number. Our solution will be to view reals in so called infinitary expansion. Under this system reals[2] correspond to infinite sequences of nonnegative integers and this correspondence is unique—every infinite sequence corresponds to a unique real and every real has corresponding to it a unique sequence. How do we find such infinitary expansions? Quite simply, as follows: suppose r is a given real. Then as in Section 2, the decimal expansion of r is found by dividing $[0, 1]$ into ten successive subintervals and seeing which subinterval r lies in, dividing this subinterval into ten successive subsubintervals and seeing which of these r lies in, and so forth. In finding infinitary expansions we do the same but here each time dividing intervals into infinitely many subintervals, a typical interval $[a, b]$ being divided into $[a, (a+b)/2)$, $[(a+b)/2, (3/4)(a+b))$, $[(3/4)(a+b), (7/8)(a+b))$, $[(7/8)(a+b), (15/16)(a+b))$,.... Thus, for example, $1, 0, 2, 1$,... is the infinitary expansion of the unique real number which simultaneously lies in $[1/2, 3/4)$, $[1/2, 5/8)$, $[19/32, 39/64)$, $[79/128, 159/256)$,....

Keeping in mind this way of viewing reals as infinite sequences of nonnegative integers we proceed as follows: let W denote the collection of all infinite sequences of nonnegative integers. Given a subset A of W there exists a game G_A played between two players I and II as follows: I initiates play by writing a nonnegative integer n_1. II responds by writing a nonnegative integer n_2. I then writes a nonnegative integer n_3, II responds with n_4, I with n_5, II with n_6, I with n_7, and so forth ad infinitum. When they have "finished" the result is a real number, that is, an infinite sequence of nonnegative integers n_1, n_2, n_3, n_4, n_5,..., and the payoff of this play of G_A is that I wins iff that real n_1, n_2, n_3, n_4, n_5,... is a member of A. Thus as I and II are playing I tries to make the sequence wind up in A and II tries to keep it out. A *strategy* for such a game is simply a function from the set of finite sequences of nonnegative integers into the set of nonnegative integers, the

[2] Throughout the remainder of this paper we will restrict our attention exclusively to the real numbers in the interval $[0, 1]$.

idea being that a player using the strategy f would write $f(n_1, n_2 \cdots n_i)$ as his move following the initial play $n_1, n_2, \ldots, n_i$. A strategy is said to be a *winning strategy* for a given game G_A if a player using the strategy always wins.

It is clear that for any such game the two players can never both possess winning strategies. But does some one of them always have a winning strategy, that is, for each set A does there exist a winning strategy for one of the players of G_A? Assuming the axiom of choice, the answer is no, for if one can well-order the set of reals he can fairly easily diagonalize over all possible winning strategies and put together a set A such that no player has a winning strategy for G_A.

This situation is unfortunate for if winning strategies always did exist for such games mathematicians would have powerful tools from which many wonderful results would follow. Here is a sample:

THEOREM.[3] *Assume that for every set A of reals there exists a winning strategy for G_A. Then every set of reals is Lebesgue measurable.*

Remark. Before launching into the technicalities of proving the theorem at hand we might make some vague remarks about the argument to be used. As expected, the problem will quickly reduce to our seeking an open covering of small measure for a given set, and such a covering we shall be able to produce quite easily from a winning strategy for a particular game to be defined. The idea is that the strategy will choose a covering for each individual point in our given set *in a uniformly continuous way*: the closer together two points in our set are, the more similar their coverings will be. Because of this the resulting covering of the entire set will be extremely efficient, that is, will cover the set as closely as possible.

Proof. It is routine to see that in order to show every set of reals Lebesgue measurable it suffices to show that any set of reals with inner measure 0 has outer measure 0. This can easily be done as follows: suppose that B is a set with inner measure 0 and let $\varepsilon > 0$ be given. We wish to find an open covering of B having measure $\leqslant \varepsilon$. (Without loss of generality we can clearly assume that B is a subset of the interval $[0, 1]$). Now it is clear that one can systematically assign to each finite union of open intervals with rational endpoints a unique nonnegative integer—one way to do this might be to assign to the union of $(i_1/i_2, i_3/i_4), (i_5/i_6, i_7/i_8), \ldots, (i_k/i_{k+1}, i_{k+2}/i_{k+3})$ the number $2^{i_1}3^{i_2}5^{i_3}7^{i_4}11^{i_5} \cdots p_n^{i_n} \cdots p_{k+3}^{i_{k+3}}$, where p_n is the nth prime number. At any rate, assuming that some such coding of intervals in terms of

[3] In considering this theorem and its proof our context will, of course, be one without the axiom of choice.

numbers is specified, let us consider the infinite game whose playoff is as follows: player II wins the play $n_1, n_2, n_3, n_4, n_5, \ldots$ if both

(1) his individual moves $n_2, n_4, n_6, n_8, n_{10}, \ldots$ are the code numbers for a sequence of unions of intervals such that for each i the union of intervals coded by n_{2i} has measure less than $\varepsilon/2^{4i}$, and

(2) if player I's moves, $n_1, n_3, n_5, n_7, n_9, \ldots$, when written $.n_1 n_3 n_5 n_7 n_9 \cdots$ happen to constitute a *binary* expansion of a member of B, then that member of B lies in one of the unions of intervals coded by $n_2, n_4, n_6, n_8, n_{10}, \ldots$.

Now although we used lots of words to describe this game it is routine to see that for some set A contained in W, I and II are just playing G_A. Thus, by the hypothesis of our theorem, let f be a winning strategy for this game. Now if we can prove that f is actually a winning strategy for player II we would have our desired cover. For in this case, by the definition of the game, it is easy to see that if Q denotes the collection of all those unions of intervals whose code numbers appear as f is played against all possible sequences of 0's and 1's as I's moves, then Q constitutes an open covering of B of measure as most ε. Thus to complete our proof we need only show that player I can *never* have a winning strategy for the game described above. With one simple observation and one well-known fact, though, this is easy. For suppose g were a winning strategy for player I for the above game. Then we can view g as entailing a map g^* from $[0, 1]$ into $[0, 1]$ defined as follows: given a real r in $[0, 1]$ let $n_2, n_4, n_6, n_8, n_{10}, \ldots$ be r's unique infinitary expansion. Then $g^*(r)$ is defined to be the real with binary expansion $.n_1 n_3 n_5 n_7 n_9, \ldots,$ where $n_1, n_3, n_5, n_7, n_9, \ldots$ is such that $n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8, n_9, n_{10}, \ldots$ is the play of G_A where g is used against player II playing $n_2, n_4, n_6, n_8, n_{10}, \ldots$. Now it is routine to see that g^* is continuous at each irrational and hence if we denote by P the range of g^*, P is a so-called analytic set. Also, since g is assumed to be a winning strategy for player I, P is a subset of B. But as is well known, analytic sets are Lebesgue measurable and thus, as the inner measure of B is 0, P must have measure 0. So let $n_2, n_4, n_6, n_8, n_{10}, \ldots$ be a sequence of integers with the property that the unions of intervals coded by them form an open covering of P and that for any i the union of intervals coded by n_{2i} has measure at most $\varepsilon/2^{4i}$. Then if II simply plays the integers $n_2, n_4, n_6, n_8, n_{10}, \ldots$ against I's playing via the strategy g the definition of P and of $n_2, n_4, n_6, n_8, n_{10}, \ldots$ easily tells us that the resulting play is a win for II. This contradicts the assumption that g is a winning strategy for I and so our theorem follows. ∎

4

As we mentioned early in Section 3, there are sets A such that neither player has a winning strategy for G_A. Let us call a set A of reals[4] such that some player DOES have a winning strategy for G_A *determinate.* As we hope the theorem just proved begins to indicate, the question of which sets of reals are determinate is a mathematically important one. Indeed, the idea of using winning strategies for various games to "continuously uniformize" sets of reals has been applied with great success to mathematical contexts besides measure theory, often with striking results, and it is likely that such techniques have yet to fully develop.

Well, then, which sets of reals are determinate? As an initial result, it is fairly easy to see that open sets are:

THEOREM. *If A is an open set of reals then A is determinate.*

Proof. Suppose that A is open. We wish to find a winning strategy for some player of G_A. The key property associated with A's being open, and that upon which our proof hinges, is the following: if a real $n_1, n_2, n_3, n_4, n_5,\ldots$ is a member of A then it is forced to be so as a result of some initial segment, that is, for some initial segment $n_1, n_2,\ldots, n_i$ of the real *any* real which begins with $n_1, n_2,\ldots, n_i$ is a member of A. At any rate, we proceed as follows: suppose that player I does not have a winning strategy for G_A. Then consider the strategy for II whereby at each of his moves II writes the least natural number so that from I's very next move on he still has no winning strategy. That is, after I's first move n_1 II writes the least integer n_2 so that from I's very next move on he still has no winning strategy. (If it is impossible for II to make such a move here then I clearly had a winning strategy from the very beginning of the game, contradicting our initial assumption). I then writes n_3 and II responds with n_4, the least nonnegative integer such that from move 5 on I still has no winning strategy. And so forth. Now although we described this II strategy with lots of words it clearly corresponds to a formal strategy as defined earlier, *and we claim that this is in fact a winning strategy for II for G_A*. For suppose $n_1, n_2, n_3, n_4, n_5,\ldots$ is a play of G_A with II using this strategy in which he lost. Then since A is open there is an integer i such that any real beginning $n_1, n_2,\ldots, n_i$ is a member of A. But this implies that when it came to I's very next move of the play of this game after the ith he *did* have a winning strategy—he could, in fact, have done anything and still won. This contradicts the fact that II played in such a way that at every initial stage of

[4] Recall that we are thinking of infinite sequences of nonnegative integers as reals (and vice versa) through the notion of infinitary expansion.

play I never had a winning strategy to complete play. The theorem is thus established. ▮

The remainder of our paper will basically be devoted to examining the question of which sets are determinate. As it turns out, things become difficult once we look at sets more elaborate than open sets. For example, it was quite a while after open sets were observed to be determinate that countable intersections of open sets (so-called G_δ sets) were proved determinate, and the another good number of years before countable unions of G_δ's (so called $G_{\delta\sigma}$'s) were shown to be determinate. A natural question to ask is "Is every Borel set of reals determinate?" This had been one of the main open problems in set theory for many years and was just recently resolved by Martin in the affirmative.

What we shall eventually present in the next section is a proof of a stronger fact namely that every analytic set is determinate.[5] Unlike Martin's recent proof of Borel determinateness, however, the proof we will use here (also due to Martin) will require stronger assumptions than just the axioms of Zermelo–Fraenkel set theory. This is where so called "large cardinals" come into play.

The argument to be used in proving analytic sets determinate, though not unusually long, will be fairly subtle and so, in the way of motivation, let us introduce the basic idea by proving a simpler result, namely, that every countable intersection of open sets is determinate. This will not be the original proof of G_δ determinateness, but because of the simplicity of G_δ's relative to analytic sets we will still be able to carry out our proof *entirely within ZFC*. This is because only an *outright provable* modification of the large cardinal axiom is needed to make the method work here.

THEOREM. *Every countable intersection of open sets is determinate.*

Proof. Suppose that B is a given countable intersection of open sets. For definiteness let us take $B = \bigcap_{i=r}^{\infty} A_i$, where each A_i is open. Now let us consider two players engaged in playing G_B. As play proceeds, the real being produced may become "captured" by various A_i, for membership of reals in open sets is determined entirely by initial segments. A systematic way to keep track of this might be as follows: as play unfolds we shall begin listing those A_i which are already known to contain the real being produced. However, to keep things orderly we will not list any A_i unless every A_j for $j < i, j \geqslant 1$ has already wound up on our list, and we shall list at most one A_i

[5] Recall that a set of reals in analytic if it is equal to the range of a continuous function on a closed set—it is a theorem that every Borel set is analytic and that there exist many analytic sets which are not Borel.

at every stage of play of the game. The following schematic might make our listing process clearer:

$$\begin{array}{l} n_1 n_2 n_3 n_4 n_5 n_6 n_7 n_8 n_9 n_{10} n_{11} n_{12} n_{13} n_{14} n_{15} n_{16} n_{17} n_{18} n_{19} n_{20}. \\ \qquad\; A_1 \qquad A_2 A_3 \;\; A_4 \qquad\qquad\qquad A_5 \end{array}$$

What we mean to represent here is that at stage 3 of play the real has first been captured by A_1. (That is, any real beginning n_1, n_2, n_3 is in A_1 but not so for reals beginning n_1 or for reals beginning n_1, n_2), that at stage 6 the real has first been captured by A_2, that by stage 7 the real has been captured by A_3, and so forth. Note that the real may have been captured by A_3 at some stage earlier than stage 7 but that since A_2 did not capture the real until stage 6 we had to wait until at least stage 7 to list A_3. Now it is relatively clear that we have set up the following representation theorem for G_B: a play $n_1, n_2, n, n, n_5, \ldots$ is *not* in B if and only if the listing procedure as described above terminates at some initial stage of play. For the first A_i which fails to get listed in this way fails to capture the real at any initial stage of play and hence, as that A_i is open, it fails to contain the real at all.

It is on this representation theorem that our proof of G_δ determinateness hinges. The plan is to set up an auxillary game in which player I makes moves as before but in which player II must, besides making this integer moves, make special additional moves in an attempt to "prove" that the list of A_i being formed as associated with the integer moves terminates at a finite initial stage of play. It is quite easy to make this precise. Namely, let G'_B be a game between two players, I and II, played as follows: the two players move alternately beginning with player I. Player I writes a nonnegative integer at each of his moves and player II writes a nonnegative integer at each of his. But player II also has the option at each of his moves to write, as an auxillary move, an arbitrary ordinal number. The payoff is that player II wins a given play of G'_B if and only if at each initial stage of play at which a new A_i gets listed (as associated with the integer moves to that point), player II wrote, at his next available move, an auxillary ordinal which was *smaller* than any auxillary ordinary he previously wrote. Thus an initial portion of play of G'_B might look like this:

$$\begin{array}{l} n_1 n_2 n_3 \;\; n_4 \;\; n_5 \;\; n_6 \quad n_7 n_8 \quad n_9 n_{10} n_{11} n_{12} n_{13} n_{14} n_{15} \cdots \\ \qquad\qquad \omega+\omega \quad \omega+2 \;\; \omega+1 \qquad 5 \qquad\qquad\quad 6 \\ \qquad\quad A_1 \qquad\qquad A_2 \;\;\; A_3 \qquad A_4 \qquad\qquad\qquad A_5 \end{array}$$

By inspection, II has already lost this play of the game for at his 7th move the ordinal he wrote, 6, was too large—since the list of A_i has just grown he would have had to write an ordinal smaller than his previous auxillary move. This was a 5. If he had written a 4 here he would still be in the game but if the list should grow again at least five times there would be no way for him to now win. Perhaps II should have played an ω instead of a 5 at his fifth move. But this might have only delayed his loss and not prevented it. For since the collection of all ordinals forms a *well*-ordering there does not exist an infinitely descending progression of ordinals, that is, a set of ordinals $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \ldots$ such that $\alpha_1 > \alpha_2 > \alpha_3 > \alpha_4 > \alpha_5 \cdots$—such a set of ordinals would not have a least member.

Now G'_B is an "open" game, that is, any time player I wins a particular play of G'_B, he had already won on the basis of some finite initial stage of that play. This is clear. Thus by the exact same proof used to show open sets of reals determinate we may conclude that one of the players has a winning strategy for G'_B. Of course, strategies here look somewhat different from before. Functions representing strategies for player I have as their domain not the set of finite sequences of nonnegative integers but rather the cartesian product of the set of finite sequences of nonnegative integers and the set of finite sequences of ordinals, and strategies for player II for this auxillary game are really represented by a pair of functions, the first function presenting the standard integer moves and the second function the auxillary ordinal moves.

At any rate, either player I or player II has a winning strategy for G'_B. We shall complete our proof of this theorem by showing that if player I has a winning strategy for G'_B then he has one for G_B and if player II has a winning strategy for G'_B then he has one for G_B.

Now it is essentially immediate that if player II has a winning strategy for G'_B then he has one for G_B. For if (f, g) is the pair of functions representing a winning strategy for II for G'_B, the function representing integer moves, f, is itself a winning strategy for II for G_B. Indeed, if II plays according to f in a game of G_B the associated list of A_i must be finite—if one bothered to play the moves of g off to the side, he would have written a smaller ordinal every time the list grew. The ordinals are well-ordered and so the list could not have grown forever.

The game G'_B was designed to somewhat hinder player II and so it is not surprising that if player II has a winning strategy for G'_B he has one for G_B. What is surprising is what we now prove namely that if player I has a winning strategy for G'_B then he has one for G_B. Indeed, player I's being able to win G'_B simply means that he can make integer moves in such a way as to ruin any *particular* attempt of player II at decreasing ordinals as for long as the associated list of A_i grows—it does not mean that the list of A_i associated with the integer moves would eventually grow to be infinite.

So suppose F is a winning strategy for player I for G'_B. We would like to somehow convert F into a winning strategy for player I for G_B. How do we do it? For a start, the arguments of F are pairs consisting of a finite sequence of nonnegative integers and a finite sequence of ordinals, whereas a strategy for I for G_B would be a function whose arguments were just finite sequences of nonnegative integers. Thus if I is to use F somehow in playing G_B, he would have to make *guesses* as to the auxillary ordinal moves to supply F with. This is where the problem arises. For if a game were played in this way with I guessing at ordinals the resulting sequences of nonnegative integers, $n_1, n_2, n_3, n_4, n_5, \ldots$ would not necessarily be a win for I for G_B, for the associated list of A_i would not necessarily be infinite. All we know is that the original moves *guessed* by *I* and supplied by him to F did not continue to decrease in size for as long as the list of A_i grew. The associated list of A_i for $n_1, n_2, n_3, n, n_5, \ldots,$ though, may have in fact been finite. Schematically this may look as follows:

Play of $G_B \to n_1\, n_2\, n_3\;\; n_4\;\; n_5\, n_6\, n_7\, n_8\, n_9\, n_{10}\, n_{11}\, n_{12}\, n_{13}\, n_{14}\, n_{15} \cdots$

I's ordinal guesses→ $\omega+\omega$ 15 1 0

Sequence of A_i associated with play of G_B $\Big\}$ A_1 $A_2 A_3$ A_4 $A_5 \rightsquigarrow$ no more A_i

Now as we see, the associated list of A_i is finite—since F was a winning strategy for I for the auxillary game, F just had to play out integers in such a way that the associated list of A_i grew for a longer period of time then the ordinals fed to it continued decreasing. Now some might say that I did not make very good guesses for the ordinals, that they simply ran down too fast—what if he had originally guessed and applied to F at the appropriate moments the sequence $\omega + \omega$, $\omega + 10$, $\omega + 5$, 250, 150. Then *if everything else remained the same* (*i.e., if F still played out the same integers against this new revised sequence*) the result,

$$n_1 n_2 n_3\; n_4 \quad n_5 \quad n_6 \quad n_7 \quad n_8 \quad n_9 \quad n_{10} n_{11} n_{12} n_{13} n_{14} n_{15}, \ldots$$

$$\omega+\omega \quad \omega+10 \quad \omega+5 \quad 250 \qquad\qquad 150$$

$$A_1 \qquad\qquad A_2 \quad A_3 \qquad A_4 \qquad\qquad A_5$$

would represent a play of the auxillary game in which I used his winning strategy F and lost. This would be a contradiction. Now recall just what we are contradicting: we had described I playing G_B using as a strategy F in conjunction with ordinal guesses, and had decided that it may not necessarily be a winning strategy for him for G_B. Our contradiction is to it not being a winning strategy *but only under the obviously special assumption that I's ordinal guesses have the special property that after play is complete he can go back and change the sequence of ordinals without altering the sequence of integers spewed out by F.*

Let us make entirely precise this property I's ordinal guesses should have in order that his using such guesses in conjunction with F yields a winning strategy for I for G_B. The following lemma is the key:

LEMMA. *For each positive integer j there exists a nonempty set C_j of finite sequences of ordinals such that the following three properties are met*:

(1) *for any j, the members of C_j are decreasing sequences of length j*

(2) *for any j, if $\langle \alpha_1,\ldots,\alpha_j\rangle$ and $\langle \beta_1,\ldots,\beta_j\rangle$ are in C_j then $F(s,\langle \alpha_1,\ldots,\alpha_j\rangle)=F(s,\langle \beta_1,\ldots,\beta_j\rangle)$ for any finite sequence s of integers whose associated list of A_i has length j.*

(3) *for any j, if $\langle \alpha_1,\ldots,\alpha_{j+1}\rangle$ is a member of C_{j+1} then $\langle \alpha_1,\ldots,\alpha_j\rangle$ is a member of C_j.*

To see that this lemma does the trick we simply argue as follows: define the strategy f for I for G_B by $f(s)=F(s,\langle \alpha_1,\ldots,\alpha_j\rangle)$ where the list of A_i associated with the finite sequence s has length j and $\langle \alpha_1,\ldots,\alpha_j\rangle$ is in C_j (by 2 of the lemma, it doesn't matter which sequence from C_j we use). It is now easy to see that f is a winning strategy for player I for G_B, for if not, let $n_1, n_2, n_3, n_4, n_5,\ldots$ be a play of G_B, where I uses f and loses. Then the associated sequence of A_i is finite. Suppose it has length j. Then if $\langle \alpha_1,\ldots,\alpha_j\rangle$ is a sequence in C_j it is routine from properties (1), (2), and (3) of the lemma to see that the play of the auxillary game, where I plays according to F and II plays integers as he just did in this play of G_B and plays, sequentially, the ordinals $\alpha_1, \alpha_2 \cdots \alpha_j$ at the appropriate time, is a play of G'_B, where I uses his winning strategy F and loses. This is a contradiction.

It only remains to establish the lemma, and this is where the pigeon-hole principle comes in. Indeed, the key clause in the lemma, number 2, the clause which says that the revised sequence can be plugged into F without changing F's action, is precisely what one would get from pigeon-holing. Proceeding formally, we first observe that for any set there exists a set of ordinals which cannot be mapped 1–1 into that set. This is a basic property of the ordinals numbers. Let then Q_0 be a set of ordinals which cannot be mapped 1–1 into the set of real numbers. We "prune" Q_0 to form a minimal such set, Q, as follows: if for every β in Q_0 $\{\alpha \in Q_0 \mid \alpha < \beta\}$ *can* be mapped 1–1 into the set of reals we let Q be Q_0 itself. Otherwise, where β_0 denotes the *least* ordinal β in Q_0 such that $\{\alpha \in Q_0 \mid \alpha < \beta\}$ cannot be mapped 1–1 into the set of reals, we let Q be $\{\alpha \in Q_0 \mid \alpha < \beta_0\}$. Now a subset A of Q is said to be *unbounded* in Q if for every α in Q there is a β in A such that $\alpha < \beta$. The instance of the pigeon-hole principle from which we shall derive our desired lemma is the following:

FACT. *Given any map f from Q into the set of reals, there exists an unbounded subset A of Q on which f is constant.*

Proof of fact. Let $\mathbb{R}$ denote the set of reals and suppose f maps Q into $\mathbb{R}$. If f is *not* constant on any unbounded subset of Q then $f^{-1}\{r\}$ is a bounded subset of Q for every real r. But by the "minimality" of Q, every bounded subset of Q can be mapped 1–1 into the reals and so, viewing Q as $\bigcup_{r \in \mathbb{R}} f^{-1}\{r\}$ it is easy to see that Q could be mapped 1–1 into $\mathbb{R} \times \mathbb{R}$. Since it is well known that $\mathbb{R} \times \mathbb{R}$ can be mapped 1–1 into $\mathbb{R}$ we may conclude that Q can be mapped 1–1 into $\mathbb{R}$ contradicting our definition of Q. The fact is thus proved.

The lemma now follows routinely.

Proof of lemma. We shall construct the sets C_j by induction on j. Actually, we shall find our sets C_j satisfying, in addition to (1) to (3) of the lemma, the following two properties: (4) for any j, the ordinals mentioned in the sequences in C_j are all members of Q, and (5) for any j, given any α in Q there is a sequence in C_j all of whose ordinals are larger than α.

Construction of C_1. Let $\mathbb{R}^*$ denote the collection of all functions from the set of finite sequences of nonnegative integers into the set of nonnegative integers. It is routine to check that $\mathbb{R}^*$ can be put into one to one correspondence with $\mathbb{R}$ and so we can apply our above facts about pigeon-holing now to maps from Q into $\mathbb{R}^*$ rather than maps from Q in $\mathbb{R}$. Thus if we define the map k from Q into $\mathbb{R}^*$ by "for any α in Q $k(\alpha)$ is the map in $\mathbb{R}^*$ whose value at any sequence s is $F(s, \alpha)$" (recall that F is I's strategy for the auxillary game), there exists an unbounded subset C_1 of Q on which k is constant. It is now easy to check that C_1 satisfies (1) through (5). In fact, (2) is the only nonobvious clause and it follows simply because for any α and β in C_1, since $k(\alpha) = k(\beta)$, $[k(\alpha)](s) = [k(\beta)](s)$ for any sequence s of nonnegative integers, and so $F(s, \alpha) = F(s, \alpha)$ for any sequence s of nonnegative integers.

C_k has been defined and we wish C_{k+1}: since C_k satisfies clause (5), let, for each α in Q, $\beta_1^\alpha, \beta_2^\alpha \cdots \beta_k^\alpha$ be a sequence in C_k whose least element, β_k^α, exceeds α. Then if we let the map g from Q into $\mathbb{R}^*$ be given by "for any α in Q, $g(\alpha)$ is the map in $\mathbb{R}^*$ whose value at any sequence s is $F(s, \langle \beta_1^\alpha, \ldots, \beta_k^\alpha, \alpha \rangle)$," our pigeon-holing fact applies to give us an unbounded subset C'_{k+1} of Q on which g is constant. It is now immediate that if we define C_{k+1} to be $\{\langle \beta_1^\alpha, \beta_2^\alpha, \ldots, \beta_k^\alpha, \alpha \rangle \mid \alpha \in C'_{k+1}\}$, C_{k+1} satisfies (1) through (5). By induction, the lemma follows. Our proof of G_δ-determinateness is now complete. ∎

5

The main point of the previous section was to introduce and motivate Martin's technique for proving analytic sets of reals determinate. We now actually give Martin's original proof, namely, that which uses a large cardinal assumption to show that every analytic set of reals is determinate.

It would probably be useful to restate here the outline of this section of the paper. We shall begin by paralleling the argument presented in Section 4 for proving G_δ sets determinate but here we will work with an analytic set rather than a G_δ set. All will go smoothly until the middle of the proof where it will become apparent that the trick used in Section 4 involving the pigeon-hole principle does not work for analytic sets. Here is where we will have to rely upon a large cardinal axiom. The motivational discussion associated with the axiom we will use appears further on in this section and can be read now if you wish. We complete the proof of analytic determinateness at the end of this section.

Now for the proof: suppose A is a given analytic set—we would like to replace the game G_A with an "open" auxillary game based on an appropriate representation theorem for analytic sets.

Now what sort of representation theorem for A would we like? Well, in analogy with the G_δ case, there are two features it should have. The first is that as a real $n_1, n_2, n_3, n_4, n_5, ...$ is being produced there must simultaneously be produced an associated list of objects such that that real's membership or nonmembership in A is directly related to the associated list having some property. And the second is that that property of the associated list be something which can potentially be "verified" as the list is being formed. For example, in the proof of G_δ determinateness the associated list was of open sets which had already "captured" the real being produced, the "property" of the list was that it be finite (in order for the real to not be in the G_δ at hand), and the "verification" of the list as it was being formed involved an attempt by player II to predict the list's eventual length. The idea behind "verification" of the property of the list is that if a verification of the property proves successful then sure enough the property is true of the list. And conversely, if the property turns out to be false of the list then some verification would have told us so.

At any rate, let us attempt to derive such a representation theorem for A. Well, since A is analytic there must exist a closed set C and a continuous function f such that A is the range of f on C. Thus for any real r, r is a member of A if and only if for some real t in C, $r = f(t)$. Our eventual representation will involve a search, given an r, for a t such that $r = f(t)$. And the key is that since C is closed (and hence has open complement) and f is continuous, such a search can be carried out as the real r is being built.

Let us now get down to the details. We begin by recalling that by viewing reals as infinite sequences of nonnegative integers, a real is in a given open set if and only if it "already is so" based on some initial segment, that is, if and only if there is an initial segment of that real such that *any* real with that initial segment is in the open set. Let us call such initial segments "secured" with respect to the open set. Thus since the complement of our closed set C is open, *a real t is in C if and only if every initial segment of t is unsecured with respect to the complement of C.* We next notice that by the basic δ/ε definition of continuous functions, *if g is any continuous function and p and q are reals such that $g(p)=q$ then for any initial segment s of q there is an initial segment s' of p such that any real with initial segment s' is mapped via g to a real with initial segment s.* (If s is the longest sequence of length at most that of s' such that any real with initial segment s' is mapped via g to a real with initial segment s, then we shall denote s by $g_{s'}$.) At any rate, these two facts show that as a real r is being formed, one can try building a real t in C such that $f(t)=r$. For any finite sequence s' unsecured with respect to the complement of C such that $f_{s'}$ is an initial segment of r is an initial segment of such a t. To push this a bit further, let us consider the partial ordering of the set of finite sequences of natural numbers defined by "$u < v$ if and only if f_v is a proper initial segment of f_u." Then it is immediate that for any real r, r is a member of A if and only if there is a t in C such that $f(t)=r$ if and only if *there exists an infinite collection Q of finite sequences of natural numbers linearly ordered by $<$ such that* (1) *each s in Q is unsecured with respect to the complement of C and* (2) *for each s in Q, f_s is an initial segment of r.*

This is basically our desired representation theorem, but we would like to put it into slightly neater form. So let us extend our partial ordering $<$ to a total ordering $<^*$ on the set of finite sequences of nonnegative integers defined by $s <^* s'$ if and only if $f_{s'}$ is a proper initial segment of f_s *or*, at the first place the sequences f_s and $f_{s'}$ differ, the integer at that place in $f_{s'}$ is larger. Then if we call a finite sequence s *f-r-unsecured* if s is unsecured with respect to the complement of C and f_s is an initial segment of r, we can state our desired representation theorem as follows:

LEMMA. *For any real r, r is a member of A if and only if $<^*$ is not a well-ordering of the f-r-unsecured sequences.*

Proof of lemma. If r is a member of A, let t be a member of C such that $f(t)=r$. Then clearly the set T of initial segments of the sequence t are all f-r-unsecured and yet T has no $<^*$-least member. Conversely, suppose r is a real such that some set T of f-r-unsecured sequences has no least member. Pick s_0 in T. Since there must be infinitely many members of T $<^*$-smaller than s_0, and since there are only finitely many sequences s such that at the first place the sequences s and s_0 differ the integer at that place in s is smaller, there must be an s_1 in T such that s_0 is a proper initial segment of s_1. Similarly there is an s_2 in T such that s_1 is the proper initial segment of s_2. Continuing in this way we can define an infinite sequence $s_0, s_1, s_2, s_3, s_4, s_5, \ldots$ of increasingly long members of T. If t is the limit of the s_i (that is, if t is the unique real u such that each s_i is an initial segment u) then clearly t is a member of C and $f(t)=r$. Thus r is a member of A. This completes the proof of the lemma. ▌

It is now fairly easy to set up our auxillary open game to replace G_A. For as a sequence of nonnegative integers r is being built (as, for example, when two people are engaged in playing G_A) one can begin writing the associated list L of f-r-unsecured sequences. If, when r is completely built, $<^*$ well-orders the set of sequences appearing in this associated list, r is not in A. Thus in the auxillary game player II's goal is to "prove" that the associated list of finite sequences is well-ordered by $<^*$. He will try to do this by attempting to map the associated list, as it comes out, order-preservingly into the ordinals. We can formally arrange this as follows: since there exist only countably many finite sequences of nonnegative integers let us arrange the process of writing the list associated (by our representation theorem) with a play of G_A so that at most one finite sequence is added to this list at every other stage of play. Then the auxillary game G'_A is played as follows: the two players move alternately beginning with player I. Player I writes a nonnegative integer at each of his moves and player II writes a nonnegative integer at each of his. But player II also has the option at each of his moves to write, as an auxillary move, an arbitrary ordinal number. The payoff is that player II wins a play of the auxillary game if and only if his ordinal moves constitute an order-preserving map of the associated list of f-r-unsecured sequences into the ordinals, that is, if and only if at each initial stage of play at which a new sequence gets placed in the associated list, player II wrote, at his next available move, an auxillary ordinal such that this association of ordinals with f-r-unsecured sequences preseves the $<^*$-ordering. To put this another way, if we view player II's writing an auxillary ordinal when a sequence gets listed as his mapping that sequence to the ordinal, then II wins the play of the auxillary game if and only if this mapping of finite sequences under $<^*$ into the ordinals under their usual ordering is order-preserving. It is routine to now see that this auxillary game is open since if player II fails to make his mapping order-preserving he fails at some finite stage. Furthermore, if player II wins a play of it then the list of finite sequences associated by the representation theorem with the sequence of integer moves must be well-ordered under $<^*$-after all, the usual ordering of the ordinals is a well-ordering. Thus by the representation theorem we may argue as we did in proving G_δ-determinateness to conclude that if player II has a winning strategy for the auxillary game he has one for the original game. Since the auxillary game is open, either player I or player II has a winning strategy for it, and so we will have shown A determinate if we can prove that if player I has a winning strategy for the auxillary game he has one for the original game. As was the case with G_δ-determinateness, this is the point of difficulty in the proof.

Suppose F is a winning strategy for player I for G'_A. We would like to convert F into a winning strategy for I for G_A and as before we must come up with some method for making ordinal guesses to supply to F.

The situation here, however, is more difficult. For recall that even though the ordinal guesses are essential arguments to be fed into F so as to use F in playing the original game G_A, these guesses must be such that they can be revised at a later time without changing F's moves. Whereas in the case of G_δ-determinateness it is only a *finite* sequence of ordinals we later wanted to revise, here it is not. Thus, in the G_δ case we were able to complete the proof with out straightforward definition, *by induction*, of sets C_i of finite sequences of ordinal guesses having the key property that any two sequences in a given set C_i are "indiscernible" by F. For in doing our later revisions we simply replaced an original sequence with a new one in the same C_i.

The case with analytic determinateness, however, is complicated by the fact that at the stage of argument where we would like to replace original ordinal guesses with new ones without anything else being changed (so as to have, as a contradiction, a play of G'_A, where I used F and lost), it is an infinite sequence or ordinals which must be replaced. This is the point at which we must appeal to a large cardinal axiom.

Our approach is to use a generalization of a well-known theorem of Ramsey. Ramsey's idea is as follows: by a straightforward use of the pigeon-hole principle, given any positive integer n and any partition of the collection of n-element subsets of a given infinite set χ into a finite number of pieces, there must exist infinitely many n-element subsets of χ all of which lie in the same piece—it is just a minor variation of this which we used in proving G_δ-determinateness. But what, asked Ramsey, if we seek something more than just infinitely many n-element subsets of χ all of which lie in the same piece: what if we ask for an infinite subset y of χ such that all n-element subsets of y lie in the same piece? Does there always exist such a set y for such partitions? Of course if n equals 1 the answer is obviously yes—this is just an instance of the pigeon-hole principle. But if n equals 2 or greater the situation is fairly difficult. The answer, however, remains yes. This is Ramsey's theorem.

In order to contrast the difference between Ramsey's idea and mere pigeon-holing let us introduce the following simple notation: if z is any set and n is a positive integer let us denote by $[z]^n$ the collection of n-element subsets of z. Then if we denote by ω the set of nonnegative integers, an instance of pigeon-holing is "for any map f from $[\omega]^2$ into a finite set there is an infinite subset u of $[\omega]^2$ such that f is constant on u"—an instance of Ramsey's theorem is "for any map f from $[\omega]^2$ into a finite set there exists an infinite subset v *of* ω such that f is constant on $[v]^2$."

There are various ways in which Ramsey's theorem might be extended. One such extention is the following: given, for each positive integer n, a map

f_n from $[\chi]^n$ into the set of reals, there exists an uncountable subset y of χ with the property that for each n, f_n is constant on $[y]^n$. Note that we have left open the question of what χ is. By the property stated we immediately see that χ must be at least so big that it cannot be mapped one to one into the set of reals. As it turns out, though, any set χ satisfying the above extension of Ramsey's theorem must be very much larger than this. It must be larger than the set of sets of reals, the set of sets of sets of reals, and so forth. In fact, as we shall later discuss the existence of such a set χ cannot be proven in Zermelo–Fraenkel set theory even with the axiom of choice and so we must assert its existence as a new axiom.

In the next section we shall discuss more fully the considerations associated with using such sets χ in mathematics. For the moment, however, let us assume **as a new axiom** the assertion *there exists a set χ with the following property: given, for each positive integer n, a map f_n from $[\chi]^n$ into the reals, there exists an uncountable subset y of χ such that for each n f_n is constant on $[y]^n$*[6] and finish our proof that A is determinate. Now recall that we are at the following point in our proof: we have defined an auxillary game G'_A and know that one of the players has a wnning strategy for G'_A. Furthermore we know that if player II has a winning strategy for G'_A he has one for G_A and so we are assuming that F is a winning strategy for player I for G'_A. In order to complete our argument we will convert F to a winning strategy for player I for G_A, and in analogy with the G_δ case we will do this by looking for a method for player I's making ordinal guesses.

We proceed as follows: let χ be a set of ordinals satisfying the property in our new axiom and for each n let us define a function f_n from $[\chi]^n$ into the set of reals[7] as follows: given $\{\alpha_1,\ldots,\alpha_n\}$ in $[\chi]^n$ we wish a value for $f_n(\{\alpha_1,\ldots,\alpha_n\})$. This must be a function from the set of finite sequences of nonnegative integers into the set of nonnegative integers and our definition is simply that the value of $f_n(\{\alpha_1,\ldots,\alpha_n\})$ at the finite sequence s is $F(s, \langle\alpha_{i_1},\ldots,\alpha_{i_n}\rangle)$ where $\langle\alpha_{i_1},\ldots,\alpha_{i_n}\rangle$ is just the rearrangement of $\langle\alpha_1,\ldots,\alpha_n\rangle$ such that $(s, \langle\alpha_{i_1},\ldots,\alpha_{i_n}\rangle)$ is a partial play of G'_A in which player II has not yet lost. (Note that there clearly exists such a *unique* rearrangement of $\langle\alpha_1,\ldots,\alpha_n\rangle$ provided that associated with the partial play s precisely n sequences get placed in our associated list. If s does not have this property let us *define* $f_n(\{\alpha_1,\ldots,\alpha_n\})(s)$ to be 0—we will not really be interested in this case as the proof proceeds.)

Now since χ has the property stated in our new axiom, let y be an uncount-

[6] By the axiom of choice it turns out that every set can be put into one to one correspondence with a set of ordinals and so it is easy to see that we can, without loss of generality, take our set χ to be a set of ordinals.

[7] Actually, as happened earlier, we will be using $\mathbb{R}^*$, the set of functions from the set of finite sequences of nonnegative integers into the set of nonnegative integers, instead of $\mathbb{R}$, the set of reals, here.

able subset of χ such that for each n f_n is constant on $[y]^n$. It is now fairly easy to see that we can take y to be our desired "collection of ordinal guesses." Specifically, let us define the strategy g for player I for G_A as follows: given a partial play s (where there are precisely n sequences in the list associated with s by the representation theorem) $g(s)$ is simply $F(s, \langle \alpha_{i_1}, ..., \alpha_{i_n} \rangle)$, where each α_{i_j} is a member of y and $\langle \alpha_{i_1}, ..., \alpha_{i_n} \rangle$ is a sequence such that $(s, \langle \alpha_{i_1}, ..., \alpha_{i_n} \rangle)$ is a partial play of G'_A in which player II has not yet lost. Note that since the ordinals in y are "indiscernible by F" our choice of just which α_{i_j} we use in defining $g(s)$ is irrelevant—we would get the same value for $g(s)$ whichever α_{i_j} we plugged into F. More importantly, the indiscernibility of the ordinals in y by F easily yields that this strategy g is a *winning* strategy for player I for G_A. For if not, let r be the result of a play of G_A in which player I used g and lost. Then since r is not a member of A, the list associated with r by the representation theorem, $s_1, s_2, s_3, s_4, s_5, ...$ is such that its members are well-ordered by $<^*$. But since this set of s_i under $<^*$ is a countable well-ordered set it can be mapped in an order-preserving way into any uncountable well-ordered set. In particular it can be mapped order-preservingly into y.[8] [To see this is fairly easy: Suppose $(A, <_A)$ and $(B, <_B)$ are two well-orderings, A countable and B uncountable. To see that $(A, <_A)$ can be mapped order-preservingly into $(B, <_B)$ we first show that for every p in A there is a unique order-preserving map of $\{q \in A \mid q \leqslant_A p\}$ onto an initial segment of B[9]: suppose not and let p be the $<_A$-least member of A for which this is not true. For each q in A, $q <_A p$, let h_q be the unique order-preserving map of $\{r \in A \mid r \leqslant_A q\}$ onto an initial segment of B. It is routine to check that because these maps h_q are unique, they extend one another (that is, if $r <_A q <_A p$ then h_q restricted to $\{t \mid t \leqslant_A r\}$ is just h_r). Thus we can define a map from $\{q \mid q <_A p\}$ into B by $h(q) =_{\mathrm{df}} h_q(q)$. Furthermore since A is countable it is easy to check that the range of h is a countable initial segment of B. Since B is uncountable let u denote the $<_B$-least member of B not in the range of h. Then it is routine to check that the map h' of $\{q \mid q \leqslant_A p\}$ into B given by

$$\begin{aligned} h'(q) &= h(q) \qquad \text{if} \quad q <_A p \\ &= u \qquad\quad\ \ \text{if} \quad q = p \end{aligned}$$

is the unique order-preserving map of $\{q \mid q \leqslant_A p\}$ onto an initial segment of B. This is a contradiction. We can now immediately see how to map A order-preservingly into B. For if, for each p in A, h_p is the unique map from

[8] The following 21 lines constitute a proof that any countable well-ordered set can be mapped order-preservingly into any uncountable one. Some readers may wish to temporarily skip this lemma and proceed with the next paragraph.

[9] An initial segment of a well-ordering $(W, <)$ is a subset Q of W such that whenever u is in W and $v < u$, v is in W.

$\{q \mid q \leqslant_A p\}$ onto an initial segment of B then as before, h defined by $h(p) = h_p(p)$ is an order-preserving map of A into B. ∎

At any rate, since this set of s_i under $<^*$ can be mapped order-preservingly into γ let $s_i \mapsto \beta_i$ represent one such map. Then consider the play of G'_A, where I and II make integer moves and form r and II, each time a finite sequence s_i gets placed in the associated list writes the ordinal β_i. This is easily seen to be a play of G'_A in which I played his winning strategy F and lost. From this contradiction the theorem follows. ∎

6

We would now like to spend some time discussing Ramsey's idea and related partition properties such as that used in proving analytic sets determinate.

Recall that Ramsey's theorem states that given an infinite set A, a positive integer n, and a partition of $[A]^n$ into a finite number of pieces there exists an infinite subset B of A such that $[B]^n$ is entirely contained within one of the pieces of the partition. This seemingly mysterious generalization of the pigeon-hole principle is really quite natural when viewed slightly differently: given an ordered set $(A, <)$ and an n-place property $P(\chi_1, \ldots, \chi_n)$ which applies to members of A, let us call a subset B of A a set of *order-indiscernibles* for P if given any two sequences $b_1 < b_2 < \cdots < b_n$ and $c_1 < c_2 < \cdots < c_n$ of members of B, P is true of $(b_1, \ldots, b_n)$ if and only if P is true of $(c_1, \ldots, c_n)$. Then an immediate consequence of the pigeon-hole principle is simply that given any infinite ordered set and any 1-place property there exists an infinite set of order-indiscernibles for that property. Ramsey's theorem is simply the same thing but for arbitrary properties instead of just 1-place properties. This is the context in which Ramsey discovered his theorem.

It is not at all difficult to prove Ramsey's theorem. We proceed by proving by induction on n that given any positive integer n, any infinite set A, and any partiton F of $[A]^n$ into a finite number of pieces there exists an infinite subset C of A such that $[C]^n$ is entirely contained within just one of the pieces of the partition (such a set C is said to be *homogeneous* for the partition F): If $n = 1$ we simply have an instance of the pigeon-hole principle and so this case is immediate. Now suppose as our inductive hypothesis that the theorem is true for $n = k$ and that for some infinite set A F is a partition of $[A]^{k+1}$ into m-many pieces. Let us view F as a map from $[A]^{k+1}$ into $\{1, 2, \ldots, m\}$. We wish to find an infinite homogeneous set for F, that is, an infinite subset C of A such that F is constant on $[C]^{k+1}$. Our method will be to use the inductive hypothesis by considering partitions F_p, for p in A, defined as follows: given a member p of A F_p maps $[A - \{p\}]^k$ into

$\{1, 2, \ldots, m\}$ by $F_p(\chi) =_{\mathrm{df}} F(\chi \cup \{p\})$. We shall use homogeneous sets for various F_p in finding our desired homogeneous set for F. Proceeding formally, we define by induction a sequence $p_0, p_1, p_2, p_3, p_4, p_5, \ldots$ of elements of A and a sequence $C_1, C_2, C_3, C_4, C_5, \ldots$ of subsets of A as follows: p_0 is any member of A and C_1 is any subset of A homogeneous for F_{p_0}. Once p_{j-1} and C_j have been defined we let p_j be any member of C_j and let C_{j+1} be a set homogeneous for the partition F_{p_j} restricted to $[C_j - \{p_j\}]^k$. The result of this construction is that for any p_{j_0} in our sequence, $\{p_i | i > j_0\}$ is homogeneous for $F_{p_{j0}}$. Thus let us call p_{j0} an l-point if the range of $F_{p_{j0}}$ on $[\{p_i \mid i > j_0\}]^k$ is l. By the pigeon-hole principle there is some l_0 between 1 and m such that infinitely many p_j's are all l_0-points. If C denotes this set of l_0-points it is easy to check that C is our desired infinite set homogeneous for F. This completes our proof of Ramsey's theorem. ∎

Now what about the possibility of proving various extensions of Ramsey's theorem? For example, what about the partition property used to complete our proof of analytic determinateness? As we examine this problem we immediately see that there are three directions in which we would like to extend. First of all, although our proof of Ramsey's theorem produces *countable* homogeneous sets, our proof of analytic determinateness required an uncountable homogeneous set. Secondly, Ramsey's theorem comments on partitions into finitely many pieces whereas our proof of analytic determinatenss involved partitions into uncountably many pieces. And finally, in the property used in proving analytic determinateness a single set homogeneous simultaneous for infinitely many given partitions is used. Can we routinely find such sets?

As it turns out the first two forms of generalization are possible in ZFC and the third is not. Indeed, by a simple (though subtle) modification of Ramsey's original argument it is easy to prove that given a positive integer n and a set z there exists a set u "so large" that given any partition F mapping $[u]^n$ into z there exists an uncountable subset v of u homogeneous for F. But by various techniques of set theory, one can show that ZFC is not strong enough to prove the existence of a set u with the property that for any countable collection of partitions of the form $F\colon [u]^n \to z$ there exists a single infinite set homogeneous for all of them. Any such set χ turns out to be larger than anything definable with just the axioms of ZFC. Even for the case in which z is a set with just 2 elements the associated χ must be larger than the set of reals, the set of sets of reals, the set of sets of sets of reals, and so forth ad infinitum. Perhaps more to the point, is that the existence of such a set χ implies that ZFC is consistent, and thus by the well-known theorem of Gödel that no sufficiently rich consistent axiomatic theory can prove its own consistency, we must conclude that no set with this property can be proved to exist in ZFC.

So where does this leave us with respect to proving theorems by assuming the existence of such sets χ? Must they remain corollaries of some relatively obscure nonprovable partition property? Not necessarily. For a start, the property used in proving analytic determinateness is itself provable from a quite well-known set theoretic axiom, namely, that asserting the existence of a measurable cardinal. But more important is that the basic method used in proving analytic determinateness can be modified so as to eliminate dependence upon a nonprovable partition property. An example of this appears in our proof of G_δ determinateness given in Section 4. Quite impressively, such a large cardinal inspired proof provided the first correct argument *in ZFC* for the determinateness of $G_{\delta\sigma\delta}$ sets (countable intersections of countable unions of G_δ's). This theorem had been a key open question for many years and it is a good bet that it would never have been proved as early as it was were it not for Martin's original proof, using a large cardinal, of analytic determinateness. (Note: It turns out that analytic determinateness itself is not provable within ZFC.)

So far as a general study of the different large cardinal axioms and the relationships between them is concerned, the theory is extensive indeed. Some of these axioms can be stated in terms of partition properties and some in terms of the existence of various types of measures. Others must be stated in terms of entirely different concepts. And quite remarkably, these seemingly independent axioms ae not independent at all. Some directly imply others, some have the same consequences as others, and so forth. And after a fairly extensive look at large cardinal theory one begins to feel that such coincidence provides compelling evidence for the "correctness" of these axioms—the isolated "reasonable" guesses at ways to extend ZFC have turned out to cohere dramatically.

But at the moment, the mathematically most interesting aspect of large cardinal theory lies not in the details of the theory itself nor in the fact that this theory seems to be "correctly" extending mathematics. Rather it lies in the applications to standard mathematics inspired by large cardinal axioms. A case in point, as indicated earlier, is the technique we presented in showing analytic sets of reals determinate. And hopefully, large-cardinal-inspired techniques will be applied to more conventional mathematics as well. The potential is certainly present and ought definitely to be explored.

CREDITS

The theorem appearing in Section 3 is due to Mycielski and Swierczkowski. The proof of it which we present is due to Harrington. The first theorem appearing in Section 4 is due to Gale and Stewart. The second theorem appearing in Section 4 is due to Wolfe (although the proof given is not). The generalization of Ramsey's theorem mentioned in Section 6 are due to Erdös and Rado. The first proof in ZFC of $G_{\delta\sigma\delta}$-determinateness was given by Paris.

References

1. P. J. Cohen, The independence of the continuum hypothesis, I, II, *Proc. Nat. Acad. Sci. U.S.A.* **50** (1963), 1143–1148, **51** (1964), 105–110.
2. P. Erdös and R. Rado, A partition calculus in set theory, *Bull. Amer. Math. Soc.* **62**, 427–489.
3. K. Godël, "Ueber formal unentscheidbare Sätze der Principia Mathematica und verwandter Systemme I," Monatsh. Math. Phys., Vol. 38, pp. 173–198.
4. E. M. Kleinberg, Infinitary combinatorics, *in* "Proceedings, Cambridge Summer School in Mathematical Logic 1971," Lecture Notes in Math. No. 337, pp. 361–418, Springer-Verlag, Berlin/New York/Heidelberg, 1973.
5. D. A. Martin, Measurable cardinals and analytic games, *Fund. Math.* **66** (1969–1970), 287–291.
6. D. A. Martin, Borel determinacy, *Ann. of Math.* **102** (1975), 363–371.
7. J. Mycielski and S. Swierczkowski, On the Lebesgue measurability and the axiom of determinateness, *Fund. Math.* **54**, 67–71.
8. J. B. Paris, $ZF \vdash \Sigma_4^0$ determinateness, *J. Symbolic Logic* **37** (1972), 661–667.